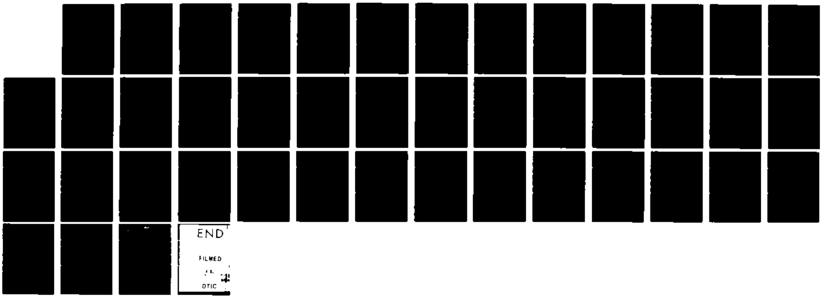
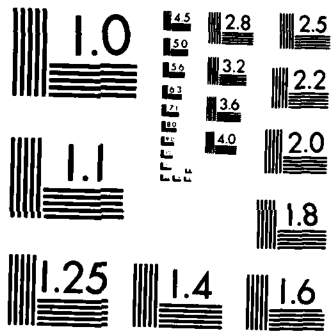
AD-A125 723 PERFORMANCE BOUNDS IN ROBUST FILTERING AND SMOOTHING (U) CONNECTICUT UNIV STORRS DEPT OF ELECTRICAL ENGINEERING AND CO.. P PAPANTONI-KAZAKOS 20 OCT 82 1/1

UNCLASSIFIED TR-82-6 AFOSR-TR-83-0110 AFOSR-78-3695 F/G 12/1 NL





MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

AFOSR-TR- 83-0110

(5)

The University of Connecticut

SCHOOL OF ENGINEERING

Storrs, Connecticut 06268

AD A125722

PERFORMANCE BOUNDS IN ROBUST FILTERING AND SMOOTHING

P. Papantoni-Kazakos

Technical Report EECS TR-82-6

October 20, 1982

A

Department of

Electrical Engineering and Computer Science

PERFORMANCE BOUNDS IN ROBUST

FILTERING AND SMOOTHING

P. Papantoni-Kazakos

Technical Report EECS TR-82-6

October 20, 1982

Performance Bounds in Robust Filtering and Smoothing

P. Papantoni-Kazakos
Department of Electrical Engineering and
Computer Science
The University of Connecticut
Storrs, CT. 06268

Abstract

Filtering and smoothing within a convex and closed family of stationary processes is considered. The perfromance criterion adopted is the mean square error. Using a saddle point game approach, we develop two classes of lower bounds for this error. Those bounds can be used as evaluation measures in the design of robust filters and smoothers.

This research was supported by the Air Force Office of Scientific Research under the grant AFOSR-78-3695C.

1. Introduction

The filtering and smoothing problems have enjoyed attention for a large number of years. The bulk of the work concentrates around the parametric model [2]. The assumption there is that the measures which generate the data and noise processes are well known. Also, the term filtering usually refers to the extraction of information data from noisy observations, while the term smoothing usually refers to the removal of outliers strictly from information data. Unfortunately, the two terms have been used sometimes inter-changeably in the literature. Finally, within the framework of the parametric model, the bulk of the work has focused on linear filters and smoothers, due to the mathematical tractability of the problem then.

The last seven years considerable attention has been given to the robust filtering and smoothing problems. The assumption there is that either the process that generates the information data, or the noise process, or both are statistically contaminated. That is, the stochastic measures that generate the processes are only members of a whole class of measures, rather than being well known. Tukey [3] introduced a variety of empirical robust smoothers, among which the "L-smoothers" (smoothers based on moving order statistics) prevail. Velleman [9] studied numerically the performance of the L-smoothers, for information processes consisting of a pure sinusoidal and additive noise with various outliers. Another class of robust smoothers is based on robustified splines, where the basic reference on this approach is provided by Huber [1]. The smoothers in [1], [3], and [9] are mostly valuable in exploratory data analysis, where conclusions on nonobvious structural properties of the information process are sought. On the other hand, Martin considered smoothers whose objective is either the generation of outlier-free time series for fitting autoregressive-moving-average models [5], [6], or the estimation of spectral density functions [14]. Then, in contrast to the methods in [1], [3] and [9], more detailed knowledge on the statistical behavior of the smoothers is

needed. The smoothers proposed in [1], [3], [9], [5], [6], and [14] are in general nonlinear. A nice review on these smoothers, as well as an additional class of such smoothers and interpolators is provided by Martin [7]. Using the terms smoothing and filtering in their usual context (as explained in the beginning of the introduction), we can refer to studies on the asymptotic robust behavior of some linear smoothers and filters. Stuck [8] studied the minimum error dispersion in linear filtering, within the class of symmetric stable processes. Hosoya [13] considered predictive linear smoothing within the convex class of processes that is modeled by linear contamination of a well known nominal process. He formalized the problem as a saddle point game in the frequency domain, and found the saddle point solution in terms of spectral densities. For some work on robustification of Wiener filtering see [16].

Despite the work on robust smoothing and filtering mentioned above, a general theory concerning the problem is still lacking. This lack makes the comparison of different smoothing and filtering schemes virtually impossible. A first effort towards that direction, for nonlinear smoothers, is provided by Mallows [15], and it is further carried on by Martin [7]. Mallows' work, however, is directed towards the analysis of robust smoothers rather than addressing the issue of how to design such smoothers. The design issue for both smoothing and filtering is perhaps better addressed if the qualitative aspects of robustness are first carefully considered. Indeed, through the appropriate selection of performance measures, the theory of qualitative robustness provides sufficient conditions that can be used as design guidelines. For the class of memoryless processes, these sufficient conditions were provided by Hampel and they can be found in [1]. For the class of stationary processes with memory, the formalization of qualitative robustness and the subsequent sufficient conditions can be found in [10]. A first qualitative formalization of the filtering problem is presented in [12]. There, the approach used in [10] in conjuction with results from [11] are used and further extended, to provide

sufficient conditions for robust filtering. As in robust parameter estimation, the results in [12] declare that linear filtering or smoothing may be highly nonrobust. On the other hand, an appropriate selection of a memoryless nonlinearity followed by a linear filter may be robust. (Such a selection is presented in [12] for filtering in additive, Gaussian, and memoryless noise). Therefore, Hosoya's approach [13] may be nonrobust for some class of processes, if some nonlinear transformation of these processes does not preceed the linear filtering. That is, for small deviations within the class of the processes, (the measure rather than the spectrum) unreasonably large performance deviations may appear. In addition, as compared to the performance of a linear filter or smoother, the performance induced by an appropriate nonlinear such selection may be superior for all processes in the considered class. Indeed, given a well known process, the matched linear filter or smoother is, in general, the worst in terms of performance.

In the present document, we consider the filtering problem in a generalized fashion. Specifically, we consider the problem of extracting the data generated by a stationary information process with memory, from noisy observations, in general. We include the case of no noise in our general formalization; thus we include the smoothing problem. We assume that the information and noise processes (the latter may be absent) are jointly stationary, and we adopt the mean square performance criterion. We assume that the noise process (if present) is well known, and we allow the information process to vary within a prespecified convex and closed family of stationary processes. We consider the class of filters or smoothers that consist of a stationary (in general nonlinear) operation on observation data within a sliding window of fixed and finite length, followed by a linear transformation. We first show that for every fixed nonlinear operation, there exists a saddle game formalization and a unique saddle point solution. Then, we develop two classes of lower bounds on the saddle point error induced by every choice of the nonlinear operation. We use the saddle point error induced by the class of linear filters or smoothers

as an upper bound to the saddle point error induced by all possible nonlinear operations. The saddle point error performance of any given nonlinear operation is then evaluated against the developed lower bounds and the above upper bound.

2. The Model-Notation

Let $[\mu,R,X]$ denote a stationary process, with measure μ, name X, and R the real line representing, in general, the alphabet of the process. Let $[R,\nu,R]$ represent a stationary channel with input and output alphabets represented, in general, by the real line. Given μ and ν as above, we denote by $[\mu\nu^{-1},R,Z]$ the stationary process induced by the stationary process $[\mu,R,X]$ and the stationary channel $[R,\nu,R]$. In the filtering and smoothing problems, $[\mu,R,X]$ represents the information process whose data sequences must be extracted. The channel $[R,\nu,R]$ represents the transmission or observation noise and in the smoothing problem it is deterministic. The process $[\mu\nu^{-1},R,Z]$ represents the observation process in both the filtering and smoothing problems. Any filter or smoother operates then on data sequences generated by the process $[\mu\nu^{-1},R,Z]$. We will denote by X_i, Z_i the random variables representing the ith datum from the information and observation processes respectively. We will denote by x_i, z_i specific values of the random variables X_i, Z_i respectively. We will also denote:

$$\begin{aligned} X_i^j &= \{X_i, X_{i+1},\dots,X_j\} \ ; \ j \geq i \\ Z_i^j &= \{Z_i, Z_{i+1},\dots,Z_j\} \ ; \ j \geq i \\ x_i^j &= \{x_i, x_{i+1},\dots,x_j\} \ ; \ j \geq i \\ z_i^j &= \{z_i, z_{i+1},\dots,z_j\} \ ; \ j \geq i \end{aligned} \tag{1}$$

We assume that the measure μ and the measures induced by the stationary channel $[R,\nu,R]$ are absolutely continuous. Then, we denote by $f_{\mu,\nu}(z^{j-i})$ the (j-i)-dimensional density function induced by the process $[\mu\nu^{-1},R,Z]$ at the point z_i^j in the (j-i)-

dimensional Euclidean space. We will use the symbol E for expected value. Then, $E_\mu X_0^2$ will denote the expected value of X_0^2 at the measure μ, and $E_{\mu,\nu}\left\{X_0 \middle| Z_i^j\right\}$ will denote the conditional expectation of the random variable X_0 conditioned on the observation sequence Z_i^j, as induced by the process $[\mu,R,X]$ and the channel $[R,\nu,R]$. For convenience in notation, we will denote:

$$\alpha_{\mu,\nu,(i,j)}(z^{j-i}) = E_{\mu,\nu}\left\{X_0 \middle| Z_i^j = z_i^j\right\} \tag{2}$$

Let us consider a sliding block window of length ℓ, operating on observation sequences. Let k denote the number of sliding steps per shift of the sliding block window, where $1 \leq k \leq \ell$. If $k = 1$, there is an one datum sliding per unique shift of the sliding window. If $k = \ell$, each unique shift of the sliding window corresponds to a shift by a whole ℓ-size data block. This last case corresponds to a block, rather than sliding block operation. For given k and ℓ, we will denote:

$$\beta_{i,\mu,\nu}^{(k)}(z^\ell) = E_{\mu,\nu}\left\{X_0 \middle| Z_{ik+1-\ell}^{ik} = z^\ell\right\} \tag{3}$$

; where z^ℓ denotes a given value of an ℓ-dimensional data vector.

We will also denote then:

$$B_{\mu,\nu,(k),i,j}^T(z^\ell) = \left[\beta_{i,\mu,\nu}^{(k)}(z^\ell),\ldots,\beta_{j,\mu,\nu}^{(k)}(z^\ell)\right] \tag{4}$$

; where the expression in [4] symbolizes a row vector, and T means transpose.

As we already mentioned, the stationary channel $[R,\nu,R]$ may be deterministic. Then, there is no observation noise and the data extraction problem becomes the smoothing problem. Thus, in our general formalization, with the inclusion of the channel $[R,\nu,R]$, both the filtering and smoothing problems are included. Also, our

model includes the filtering problem for a process embedded in additive noise, with the noise and information processes being mutually independent.

In general, our objective is the design of an appropriate smoother or filter. We will consider the existence of an, in general, sliding block window on the observation sequences..., Z_{-1}, Z_0, Z_1,... . We will denote by ℓ the size of the window, and by k the number of data steps per single window sliding; where $1 \leq k \leq \ell$. We will denote by g_ℓ some stationary operation on ℓ-size data blocks. Then, we will consider filters and smoothers of the following form:

$$G_{g_\ell,k,\{a_i\}}(z) = \sum_i a_i \, g_\ell \, (z_{ik+1-\ell}^{ik}) \; ; \; 1 \leq k \leq \ell \tag{5}$$

; where z denotes an infinite sequence $\ldots,z_{-1},z_0,z_1,\ldots,$ and $\{a_i\}$ is a set of constant coefficients.

The filter and smoother form in (5) clearly represents a cascade operation of an, in general, nonlinear operation (represented by g_ℓ) and a linear filter or smoother (represented by $\{a_i\}$). If $\ell \to \infty$ and g_ℓ is allowed to move within the total class of all possible operations, the cascade operation in (5) clearly represents the totality of possible operations on the observation data. If $\ell = 1$, and g_1 any operation per single datum, the operation in (5) includes linear filters and smoothers. Therefore, the cascade operation in (5) represents (for varying k,ℓ,g_ℓ, and $\{a_i\}$) all possible filters and smoothers, with both sliding-block and block operational characteristics.

As mentioned before, the objective of the smoother or the filter is the extraction of the data from the information process $[\mu,R,X]$. Due to the stationarity of the model, we will assume without lack in generality that the operation $G_{g_\ell,k,\{a_i\}}(z)$ in (5) corresponds to a mapping of the datum X_0 from the process $[\mu,R,X]$. Then, adopting the mean square performance criterion, we will use the

following notation for the resulting mean square error:

$$e(\mu,\nu,\ k,g_{\ell},\{a_i\}) = E_{\mu,\nu}\left\{X_0 - \sum_j a_j\ g_{\ell}\ (Z^{jk}_{jk+1-\ell})\right\}^2 \tag{6}$$

For $\ell \to \infty$ the parameters k and $\{a_i\}$ become absolute. Then, we will denote the error in (6) as $e(\mu,\nu,\ g_\infty)$.

We will denote by $F(k,g_{\ell})$ the class of filters or smoothers described by (5) for given g_{ℓ}, ℓ, and k. This class is then generated by variations of the set $\{a_i\}$. We will denote by $F(k,\ell)$ the class that is described by (5) for given k and ℓ. The class $F(k,\ell)$ is generated by variations of both g_{ℓ} and $\{a_i\}$. We will denote by F the class described by (5) for $\ell \to \infty$. This last class clearly includes all possible filters and smoothers, and also:

$$\begin{aligned} &F(k,g_{\ell}) \subset F(k,\ell) \subset F\ ;\ \forall\ \ell \\ &F(k,\ell) \supseteq F(k,\ell.m)\ ;\ \forall\ \ell\ ;\ \forall\ m \geq\quad ;\ \forall\ k \end{aligned} \tag{7}$$

Let $F(L)$ denote the class $F(1,g_1)$ for $g(x) = x$. Then, due to the relationships in (7), we have:

$$F(L) \subseteq F(1,1) \supseteq F(k,\ell)\ ;\ \forall k{:}1 \leq k \leq \ell;\ \forall\ \ell > 1 \tag{8}$$

Therefore, for the error expression in (6) we have:

$$\begin{aligned} &\inf_{G_{g_{\ell},k,\{a_i\}} \in F(L)} e(\mu,\nu,k,g_{\ell},\{a_i\}) \geq \inf_{G_{g_{\ell},k,\{a_i\}} \in F(1,1)} e(\mu,\nu,1,g_1,\{a_i\}) \geq \\ &\qquad \geq \inf_{G_{g_{\ell},k,\{a_i\}} \in F} e(\mu,\nu,k,g_{\ell},\{a_i\}) = \inf_{G_{g_\infty} \in F} e(\mu,\nu,\ g_\infty)\ ;\ \forall\ \ell\ ;\ \forall\ k{:}\ 1 \leq k \leq \ell \end{aligned} \tag{9}$$

Thus, we can express the following proposition:

Proposition 1

For every given μ and ν, the class $\mathcal{F}(L)$ of linear filters and smoothers represents a worst case. Specifically, the infimum of the mean square error in (6) within this class is, in general, larger than the parallel infimum within the class $\mathcal{F}(1,1)$. The global infimum is achieved within the class $\mathcal{F}$.

Proposition 1 expresses formally what we stated informally in the introduct: That the class of linear filters and smoothers can be used to represent an upper bound on the mean square error.

In the present document, we are interested in robust smoothing and filtering. We thus consider a family M of measures μ, rather than a well-known measure μ_0. For convenience in notation, we will assume that the noise channel $[R,\nu,R]$ is well known. In our general derivations, however, this assumption is not necessary. The only assumption we need there is just the consideration of a convex and closed family of joint measures.

Given a well-known noise measure ν (that may be deterministic) and a convex and closed family M of information processes μ, we will model robustness as a saddle point game, and we will search for saddle point solutions. We will initially ignore the qualitative aspects of robustness, since we are basically searching for performance lower bounds within the classes M and $\mathcal{F}$. We will come back to qualitative robustness later.

We conclude this section by noticing that all the sets $\mathcal{F}(L)$, $\mathcal{F}(k,g_\ell)$, $\mathcal{F}(k,\ell)$, and $\mathcal{F}$ in (7) are convex, and by stating the general formalization for a saddle point game. Given two sets M and S, given a payoff function $K(\mu,\lambda)$; where $\mu\in M$ and $\lambda\in S$, the game on $M\times S$ with payoff $K(\mu,\lambda)$ has a saddle point solution if and only if there exists a pair $(\mu^*,\lambda^*) \in M\times S$ such that:

$$\forall \mu \varepsilon M \;;\; K(\mu,\lambda^*) \leq K(\mu^*,\lambda^*) \leq K(\mu^*,\lambda) \;;\; \forall\, \lambda \varepsilon S \tag{10}$$

The saddle point solution is unique if and only if the pair (μ^*,λ^*) that satisfies (10) is unique. The quantity $K(\mu^*,\lambda^*)$ is called the saddle value of the game.

3. The Game for Fixed Nonlinearity

Let g_ℓ be a fixed, well-known stationary and deterministic operation. Let ℓ be known and finite. Then, let us consider the convex class $F(k,g_\ell)$ of filters and smoothers, where $F(k,g_\ell)$ is defined in the previous section and k is assumed fixed and known. Let the noise channel $[R,\nu,R]$ be well known, and let M be a convex and closed family of information processes μ. We will search for a saddle point game formalization and solution on $M\mathrm{x}F(k,g_\ell)$, using as the payoff function the error $e(\mu,\nu,k,g_\ell,\{a_i\})$ in (6). We will first consider finite sequences Z_i^j of observation data, and then asymptotically long such sequences.

Let the smoother or filter operate on the limited length sequency $Z_{-nk+i-\ell}^{(m-n-1)k}$ of observation data. Then, the smoother or filter take the form:

$$G_{g_\ell,k,\{a_i\}}\left(z_{nk+1-\ell}^{(m-n-1)k}\right) = \sum_{i=-n}^{m-n-1} a_i\, g_\ell\left(z_{ik+1-\ell}^{ik}\right) \;;\; m \geq 1 \tag{11}$$

Define:

$$V_{g_\ell}^T(z) = \left[g_\ell\left(z_{ik+1-\ell}^{ik}\right) \;;\; -n \leq i \leq m-n-1\right] \tag{12}$$

$$R(\mu,\nu,g_\ell) = E_{\mu,\nu}\left\{V_{g_\ell}(z)\; V_{g_\ell}^T(z)\right\} \tag{13}$$

; where the expression in (12) denotes a row vector. Then, it is straightforward to obtain the following result:

$$\inf_{G_{g_\ell,k,\{a_i\}} \in F(k,g_\ell)} e(\mu,\nu,k,g_\ell,\{a_i\}) = E_{\mu,\nu} X_0^2 - E_{\mu,\nu}\left\{X_0 V^T_{g_\ell}(Z)\right\} R^{-1}(\mu,\nu,g_\ell)\, E_{\mu,\nu}\left\{X_0 V_{g_\ell}(Z)\right\} \tag{14}$$

; where Z denotes the sequence $Z^{(m-n-1)k}_{-nk+1-\ell}$.

The infimum in (14) is actually realized for:

$$\{a_i^*\} \ : \ A^*_{\mu,\nu,g_\ell} = R^{-1}(\mu,\nu,g_\ell)\, E_{\mu,\nu}\left\{X_0 V_{g_\ell}(Z)\right\} \tag{15}$$

; where A^*_{μ,ν,g_ℓ} denotes the column vector of the appropriate coefficients $\{a_i \ ; \ -n \leq i \leq m-n-1\}$.

Let us consider the game on $M \times F(k,g_\ell)$ with the payoff function $e(\mu,\nu,k,g_\ell, \{a_i \ ; \ -n \leq i \leq m-n-1\})$. We will show that the game has a saddle point solution. We will do that in two steps.

Let us denote:

$$e(\mu,\nu,k,g_\ell) = \inf_{G_{g_\ell,k,\{a_i\}} \in F(k,g_\ell)} e(\mu,\nu,k,g_\ell,\{a_i\}) \tag{16}$$

As the first step in the proof for the solution of the game, we present the following lemma:

Lemma 1

The error $e(\mu,\nu,k,g_\ell)$ in (16) is concave in M. Therefore, there exists some μ^* in M such that:

$$e(\mu^*,\nu,k,g_\ell) = \sup_{\mu \in M} e(\mu,\nu,k,g_\ell)$$

The proof of lemma 1 is in appendix A.

The measure μ^* that satisfies the supremum in lemma 1 may not be unique. It is unique, however, within a class M of measures that induces distinct expectations

$E_{\mu,\nu}\left\{X_0 \, V_{g_\ell}(Z)\right\}$ and matrices $R(\mu,\nu,g_\ell)$.

We express now the second step in the proof for the existence of a solution to the game, in the following theorem.

Theorem 1

Let us consider the game on $M\times F(k,g_\ell)$ with the payoff function $e(\mu,\nu,k, g_\ell, \{a_i ; -n \le i \le m-n-1\})$. The game has a saddle point solution. That is, there exists a pair $(\mu^*,\{a_i^*\})$ in $M\times F(k,g_\ell)$ such that:

$$\forall\mu\varepsilon M \; ; \; e(\mu,\nu,k,g_\ell,\{a_i^*\}) \le e(\mu^*,\nu,k,g_\ell,\{a_i^*\}) \le e(\mu^*,\nu,k,g_\ell,\{a_i\})$$

$$; \; \forall \; \{a_i\} \; \varepsilon \; F(k,g_\ell)$$

The pair $(\mu^*,\{a_i^*\})$ is such that:

$$e(\mu^*,\nu,k,g_\ell,\{a_i^*\}) = e(\mu^*,\nu,k,g_\ell) = \sup_{\mu\varepsilon M} \; e(\mu,\nu,k,g_\ell)$$

$$\{a_i^*\} = A^*_{\mu^*,\nu,g_\ell}$$

; where A_{μ^*,ν,g_ℓ} is given by (15).

According to the theorem, to find a folution of the game, it is sufficient to first express the infimum in (16) for every μ in M, and to then find the supremum of the resulting expression, with respect to μ in M. The proof of the theorem is in appendix A, and a unique solution exists if uniqueness in lemma 1 is satisfied.

We will now consider the asymptotic case, where the observation sequences used in the smoothing or filtering extend to the infinite past and the infinite future. That is, we will in general consider noncausal filters and smoothers. This is done for convenience here. Some causal filters of predictive nature and belonging to the general class $F(1,g_1)$ have been considered in [12].

Let the filter or smoother be allowed the following asymptotic form:

$$G_{g_\ell, k, \{a_i\}}(z) = \sum_{i=-\infty}^{\infty} a_i \, g_\ell(z_{ik+1-\ell}^{ik}) \tag{17}$$

Then, define:

$$H(w) = \sum_{k=-\infty}^{\infty} a_k \, e^{-jkw} \tag{18}$$

$$S_{\mu,\nu,g_\ell}(w) = \sum_{m=-\infty}^{\infty} e^{-jmw} \, E_{\mu,\nu}\left\{ g_\ell(Z_{1-\ell}^{o}) \; g_\ell^{*}\,(Z_{mk+1-\ell}^{mk}) \right\} \tag{19}$$

; where $E_{\mu,\nu}\left\{ g_\ell(Z_{ik+1-\ell}^{ik}) \; g_\ell^{*}\,(Z_{mk+1-\ell}^{mk}) \right\} = f(i-m)$, due to the stationarity of μ,ν, and the operation g_ℓ, and where * means conjugate.

Also, define:

$$S_{\mu,\nu,g_\ell,\beta}\,(w) = \sum_{m=-\infty}^{\infty} e^{-jmw} \, E_{\mu,\nu}\left\{ \beta_{m,\mu,\nu}^{(k)}\,(z^\ell) \; g_\ell^{*}\,(z^\ell) \right\} \tag{20}$$

;where z^ℓ denotes some ℓ-dimensional vector of consequtive data from the observation process, and $\beta_{m,\mu,\nu}^{(k)}(z^\ell)$ the conditional expectation defined in (3).

The expressions in (18), (19), and (20) clearly signify Fourier transforms, where $H(w)$ is the asymptotic representation of the set $\{a_i\}$ of coefficients. Substituting the error expression $e(\mu,\nu,k,g_\ell,\{a_i\})$ in (6) by $e(\mu,\nu,k,g_\ell,H)$ in this case, we easily find that the Fourier transform $H_{\mu,\nu,g_\ell}^{*}(w)$ corresponding to the vector A_{μ,ν,g_ℓ}^{*} in (15) is given by the following expression.

$$H_{\mu,\nu,g_\ell}^{*}(w) = \frac{S_{\mu,\nu,g_\ell\beta}(w)}{S_{\mu,\nu,g_\ell}(w)} \tag{21}$$

Also, the infimum $e(\mu,\nu,k,g_\ell)$ in (16) takes now the following form.

$$e(\mu,\nu,k,g_\ell) = E_\mu X_0^2 - \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{\left|S_{\mu,\nu,g_\ell,\beta}(w)\right|^2}{S_{\mu,\nu,g_\ell}(w)} dw \tag{22}$$

Since the length of the observation sequences used by the filter or smoother was never used in the proof of theorem 1, we can express a theorem parallel to it for the asymptotic case.

Theorem 2

The game on $\mathcal{M}\times\mathcal{F}(k,g_\ell)$ with the payoff function $e(\mu,\nu,k,g_\ell,H)$ has a saddle point solution. That is, there exists a pair (μ^*,H^*) in $\mathcal{M}\times\mathcal{F}(k,g_\ell)$ such that:

$$\forall\mu\varepsilon\mathcal{M} \; ; \; e(\mu,\nu,k,g_\ell,H^*) \leq e(\mu^*,\nu,k,g_\ell,H^*) \leq e(\mu^*,\nu,k,g_\ell,H); \; \forall \; H \; \varepsilon \; \mathcal{F}(k,g_\ell)$$

The pair (μ^*,H^*) is such that

$$e(\mu^*,\nu,k,g_\ell) = \sup_{\mu\varepsilon\mathcal{M}} e(\mu,\nu,k,g_\ell)$$

$$H^*(w) = H^*_{\mu^*,\nu,g_\ell}(w)$$

; where $e(\mu,\nu,k,g_\ell)$ and $H^*_{\mu,\nu,g_\ell}(w)$ are given by (22) and (21) respectively.

In this section, we showed that if in the cascade-type filter or smoother in (5) we fix the operation g_ℓ, we can formulate a saddle point game in $\mathcal{M}\times\mathcal{F}(k,g_\ell)$ whose solution exists. Furthermore, this solution corresponds to mean-square error minimization in the worst case. It does not necessarily guarantee performance continuity within the class $\mathcal{M}$. For such continuity the sufficient conditions of the qualitative robustness [10,11,12] should be satisfied. That can be accomplished through the appropriate selection of the operation g_ℓ, and it may also require that

only a subset of the set $F(k,g_\ell)$ be considered in the filter or smoother selection. An example of the above can be found in [12].

4. The Unrestricted Game

In this section, we will consider the case where the filter or smoother is described by (5) for $\ell \to \infty$. That is, the class F of section 2 is adopted. We will search for a game formalization and solution on MxF; where M a convex and closed family of measures μ. As in section 2, we will use the symbol g_∞ to signify filters and smoothers in F. We will use the symbol g_∞ even in the case where the filter or smoother operates on observation sequences of finite length.

Let the filter or smoother operate on the observation sequence Z_i^j. The, we will denote as in (2):

$$\alpha_{\mu,\nu,(i,j)}(z^{j-i}) \triangleq E_{\mu,\nu}\left\{X_0 \middle| Z_i^j = z_i^j\right\} \tag{23}$$

Given some measure μ in M it is then well known that:

$$\inf_{g_\infty \varepsilon F} e(\mu,\nu,g_\infty) = e(\mu,\nu,\alpha_{\mu,\nu,(i,j)}) = E_\mu X_0^2 - E_{\mu,\nu}\left\{\alpha^2_{\mu,\nu,(i,j)}(z^{j-i})\right\} \tag{24}$$

; where $e(\mu,\nu,g_\infty)$ the mean square error in (6) for $\ell \to \infty$.

Using the payoff function $e(\mu,\nu,g_\infty)$, we are searching for a saddle point solution on MxF. As in section 3, we proceed in two steps. The first step is represented by the following lemma.

Lemma 2

The error $e(\mu,\nu,\alpha_{\mu,\nu,(i,j)})$ in (24) is concave in M. It is strictly concave iff for any two μ_1,μ_2 in M the conditional expectations $\alpha_{\mu_1,\nu,(i,j)}(z^{j-i})$ and $\alpha_{\mu_2,\nu,(i,j)}(z^{j-1})$ in (23) are not equal almost everywhere on R^{j-i}.

The proof of lemma 2 is in appendix B.

We now show the existence of a saddle point solution of the game on MxF with payoff function $e(\mu,\nu,g_\infty)$, with a theorem.

Theorem 3

Consider the game on MxF with payoff function $e(\mu,\nu,g_\infty)$. The game has a saddle point solution. That is, there exists a pair (μ^*,g_∞^*) on MxF such that:

$$\forall \mu \varepsilon M \ ; \ e(\mu,\nu,g_\infty^*) \leq e(\mu^*,\nu,g_\infty^*) \leq e(\mu^*,\nu,g_\infty) \ ; \ \forall \ g_\infty \ \varepsilon \ F$$

; where $e(\mu,\nu,g_\infty)$ is given by (6) with $\ell \to \infty$.

The pair (μ^*,g_∞^*) is such that:

$$\mu^* : e(\mu^*,\nu,g_\infty^*) = \sup_{\mu \varepsilon M} e(\mu,\nu,\alpha_{\mu,\nu,(i,j)})$$

$$g_\infty^* = \alpha_{\mu^*,\nu,(i,j)}$$

If there are no μ_1,μ_2 in M such that $\alpha_{\mu_1,\nu,(i,j)}(z^{j-i})$ and $\alpha_{\mu_2,\nu,(i,j)}(z^{j-i})$ are equal almost everywhere on R^{j-i}, the pair (μ^*,g_∞^*) is also unique.

The proof of the theorem is in appendix B. Since in the proof, the dimensionality (j-i) is no where used, the statement holds also asymptotically. According to the theorem, to find the solution of the game, it is sufficient to find the minimum mean square error in F for each measure μ, and to then find the supremum of the result on M.

It is clear that since $F(k,g_\ell) \subset F$; $\forall k,\ell,g_\ell$, the error $e(\mu,\nu,k,g_\ell)$ in (16) is, in general, bounded from below by the error $e(\mu,\nu,\alpha_{\mu,\nu,(i,j)})$ in (24) for fixed Z_i^j, and for every k,ℓ,g_ℓ, and μ in M. Therefore, the supremum of $e(\mu,\nu,k,g_\ell)$ on M is also then bounded from below by the supremum of $e(\mu,\nu,\alpha_{\mu,\nu,(i,j)})$ on M. Thus, we can express the following corollary.

Corollary 1

Given a convex and closed class M of measures μ, given the sequence z_i^j of observations on which the filter or smoother operates, the saddle value $e(\mu^*,\nu,\alpha_{\mu^*,\nu,(i,j)})$ of the unrestricted game in theorem 3 bounds from below the saddle value $e(\mu^*,\nu,k,g_\ell)$ of the restricted game in theorem 1, for every k,ℓ, and g_ℓ choice. Thus, $e(\mu^*,\nu,\alpha_{\mu^*,\nu,(i,j)})$ acts as a global lower bound on the saddle value performance of all the filtering and smoothing schemes.

We will conclude this section with the presentation of the solution of the game in theorem 3, for two important classes M of information processes. We present the results in two lemmas, and we include the proof of the first lemma in appendix B.

Lemma 3

Let the noise channel $[R,\nu,R]$ be either deterministic, or Gaussian, stationary, and additive. If the latter is true, let also the process ν be independent of any information process. Let M be the convex and closed family of stationary processes with equal means and spectral densities. Then, the solution of the game in theorem 3 is provided by the Gaussian process μ^* and the corresponding linear filter or smoother $\alpha_{\mu^*,\nu}$. This is true for any dimensionality $(j-i)$ of the filter or smoother.

We observe that the saddle point solution of the game in theorem 3, within the class of measures with equal spectral densities, is satisfied by the maximum entropy measure within the class. The relationship between worst error and highest entropy seems natural. In qualitative terms, a high entropy process is less predictable. Thus, any effort for the extraction of its data should result in higher error.

In the proof of lemma 3, in appendix B, we denoted by $e_L(\mu,\nu)$ the infimum of the error in (6), on the class $F(L)$ of linear smoothers and filters, and at the measure μ. The quantity $e_L(\mu,\nu)$ is a function of the μ and ν spectral densities only, if both the measures μ and ν have zero mean and are mutually independent. Let us consider a convex and closed family S of spectral densities, and let us assume that $e_L(\mu,\nu)$ is strictly concave with respect to the spectral density of the measure μ. This is in general true for the channel model in lemma 3 (see ref. [13]). Let us consider the convex and closed family M of measures μ, such that each μ is zero mean, stationary, and has spectral density in S. Let us denote by f_μ the spectral density of the measure μ. Then, there exists a unique spectral density f^* in S, such that:

$$e_L(f^*,f_\nu) = \sup_{f_\mu \varepsilon S} e_L(f_\mu, f_\nu) \qquad (25)$$

; where f_ν is the spectral density of the measure ν, and we have represented $e_L(\mu,\nu)$ by $e_L(f_\mu, f_\nu)$ instead.

It is now clear from the approach used in the proof of lemma 3, that we can express the following lemma.

Lemma 4

Let the channel $[R,\nu,R]$ and its relationship to any information process be as in lemma 3. Let S be a convex and closed family of spectral densities, and let the error $e_L(f_\mu, f_\nu)$ be strictly concave on S sith respect to f_μ. Let M be the convex and closed family of stationary, zero mean processes, with spectral densities in S. Let the measure ν be zero mean. Then, the solution of the game in theorem 3 is provided by this Gaussian measure μ^* whose spectral density f_{μ^*} satisfies:

$$e_L(f_{\mu^*},f_\nu) = \sup_{f_\mu \varepsilon S} e_L(f_\mu, f_\nu)$$

From lemma 4, it is clear that if the unrestricted family F of smoothers and filters is considered, the solution of a saddle point game within a convex and closed family of spectral densities is only an intermediate step towards the solution of the larger game. Due to the relationships in (8), the above statement holds even if F is substituted by $F(k,\ell)$.

5. The Game in $F(k,\ell)$

To this point, we have analyzed two extreme cases. In section 3, we fixed the parameters k and ℓ as well as the nonlinearity g_ℓ and we analyzed the game in the class $F(k,g_\ell)$ of smoothers and filters in (5). In section 4, we analyzed the unrestricted game in the global class F of filters and smoothers. Here, we will consider an intermediate case. Specifically, we will assume that the length ℓ and the sliding step k of the sliding block window in (5) are fixed, and that the non-linearity g_ℓ and the set $\{a_i\}$ of coefficients may vary. Thus, we consider the class $F(k,\ell)$ of smoothers and filters. In this case, we can not always formulate a game in the strict sense. Our approach is somewhat different than the previous approaches, and we first present the gist of it.

Let us denote by N_ℓ the class of all possible operations g_ℓ. For some μ and g_ℓ, consider the infimum $e(\mu,\nu,k,g_\ell)$ in (16). For every μ, we will find some operation $g_{\ell,\mu}$ in N_ℓ and some nonnegative function $h(\mu,\nu,k,g_{\ell,\mu})$, such that:

$$0 \leq h(\mu,\nu,k,g_{\ell,\mu}) \leq e(\mu,\nu,k,g_\ell) \;\; ; \; \forall \; g_\ell \in N_\ell \tag{26}$$

We will show that $h(\mu,\nu,k,g_{\ell,\mu})$ is concave in μ; thus if M is some convex and closed family of measures μ, there exists some μ^* in M, such that:

$$h(\mu,\nu,k,g_{\ell,\mu}) \leq h(\mu^*,\nu,k,g_{\ell,\mu^*}) \;\; ; \; \forall \; \mu \in M \tag{27}$$

If the pair (μ^*,g_{ℓ,μ^*}) that satisfies (27) happens to be such that:

$$h(\mu^*,\nu,k,g_{\ell,\mu^*}) = e(\mu^*,\nu,k,g_{\ell,\mu^*}) \tag{28}$$

then it is clear from theorem 1, expression (27), lemma 1, and expression (16), that the pair $(\mu^*,g_{\ell\ \mu^*})$ is also the solution of the game on $F(k,g_{\ell,\mu^*})xM$ with payoff function $e(\mu,\nu,k,g_{\ell\ \mu^*},\{a_i\})$. That is, the pair (μ^*,g_{ℓ,μ^*}) satisfies then the following saddle game expression:

$$\forall\mu\varepsilon M\ ;\ e(\mu,\nu,k,g_{\ell,\mu^*},A^*_{\mu^*,\nu,g_{\ell,\mu^*}}) \leq e(\mu^*,\nu,k,g_{\ell,\mu^*}) \leq e(\mu^*,\nu,k,g_{\ell},\{a_i\})\ ;$$

$$;\ \forall(g_{\ell},\{a_i\})\ \varepsilon\ F(k,\ell) \tag{29}$$

; where $A^*_{\mu^*,\nu,g_{\ell\ \mu^*}}$ is given by (15).

However, the equality in (28) may not be satisfied in general. Then, the quantity $h(\mu^*,\nu,k,g_{\ell,\mu^*})$ in (27) will provide just a lower bound on the saddle value $e(\mu^*,\nu,k,g_{\ell,\mu^*})$ of the game in (29).

We will now proceed with the search for the operation $g_{\ell,\mu}$ in N_ℓ and the study of the quantity $h(\mu,\nu,k,g_{\ell,\mu})$. We will first consider finite sequences Z_i^j of observation data, and then asymptotically long such sequences.

As in section 3, let the smoother or filter operate on the limited length sequence $Z_{-nk+1-\ell}^{(m-n-1)k}$ of observation data. Then, the form of the smoother or filter is as in (11). Let $\beta^{(k)}_{i,\mu,\nu}(z^\ell)$ and $B_{\mu,\nu,(k),i,j}(z^\ell)$ be as in (3) and (4) respectively. Let us define:

$$M_{i-j}(\mu,\nu) = M_{ij}(\mu,\nu) \triangleq E_{\mu,\nu}\left\{B_{\mu,\nu,(k),-n,m-n-1}(Z_{ik+1-\ell}^{ik})\ B^{*T}_{\mu,\nu,(k),-n,m-n-1}(Z_{jk+1-\ell}^{jk})\right\} \tag{30}$$

; where $M_{ij}(\mu,\nu) = M_{i-j}(\mu,\nu)$ due to the stationarity assumed, * means conjugate, and T means transpose.

Clearly, the matrices $M_{i-j}(\mu,\nu)$ in (30) are square, and nonnegative definite. Consider the matrix $M_0(\mu,\nu)$ and define:

$A_{M_0}(\mu,\nu)$: it minimizes with respect to the vector A the error expression

$$E_{\mu,\nu}\left\{X_0 - A^T \mathcal{B}_{\mu,\nu,(k),-n,m-n-1}(z^{\ell})\right\}^2 \tag{31}$$

$$h(\mu,\nu,k,g_{\ell,\mu}) = E_\mu X_0^2 - A_{M_0}^T(\mu,\nu)M_0(\mu,\nu)A_{M_0}(\mu,\nu) \tag{32}$$

<u>Lemma 5</u>

Let μ be given, and let $e(\mu,\nu,k,g_\ell)$ be the infimum in (16).

Let

$$g_{\ell,\mu} : g_{\ell,\mu}(z^{\ell}) = A_{M_0}^T(\mu,\nu)\ \mathcal{B}_{\mu,\nu,(k),-n,m-n-1}(z^{\ell})$$

Then,

$$0 \le h(\mu,\nu,k,g_{\ell,\mu}) \le e(\mu,\nu,k,g_\ell) \ ; \ \forall\ g_\ell \in \mathcal{N}\ell$$

; where $h(\mu,\nu,k,g_{\ell,\mu})$ is given by (32)

The proof of the lemma is in appendix C. The result in it corresponds to expression (26) in the gist of the approach. We observe that $g_{\ell,\mu}(z^{\ell})$ is a linear combination of conditional expectations of the type given by (3). We now proceed with the next step in our approach, by stating another lemma whose proof can be found in appendix C.

<u>Lemma 6</u>

The error $h(\mu,\nu,k,g_{\ell,\mu})$ in (32) is strictly concave in μ, on some convex and closed family $\mathcal{M}$ of measures μ. Therefore, there exists a unique μ^* in $\mathcal{M}$ such that:

$$h(\mu^*,\nu,k,g_{\ell,\mu^*}) \ge h(\mu,\nu,k,g_{\ell,\mu}) \ ; \ \forall\ \mu\in\mathcal{M}$$

We now summarize parts of our discussion in this section in conjuction with some straightforward observations, in a corollary.

Corollary 2

Given a convex and closed class M of measures μ, given the sequence Z_i^j of observation data on which the filter or smoother operates, the supremum $h(\mu^*,\nu,k,g_{\ell,\mu^*})$ in lemma 6 bounds from below the saddle value $e(\mu^*,\nu,k,g_\ell)$ of the restricted game in theorem 1, for fixed k,ℓ, and all g_ℓ choices. Furthermore, the lower bound represented by $h(\mu^*,\nu,k,g_{\ell,\mu^*})$ is in general tighter than the lower bound given by the saddle value $e(\mu^*,\nu,\alpha_{\mu^*,\nu,(i,j)})$ of the unrestricted game in theorem 3 and corollary 1, for filters or smoothers in the class $F(k,\ell)$. The two bounds approach each other as ℓ increases, becoming asymptotically identical.

Let us now consider the asymptotic case, where the observation sequences used in the smoothing or filtering extend to infinite past and infinite future. Then, we will use the notation and the procedure used in section 3. Let us define the following Fourier transforms:

$$S_{\mu,\nu,\beta}(\lambda,x) = \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} e^{jn\lambda-jmx} E_{\mu,\nu} \left\{\beta_{n,\mu,\nu}^{(k)}(Z^\ell)\ \beta_{m,\mu,\nu}^{*(k)}(Z^\ell)\right\} \tag{33}$$

$$S_{\mu,\nu,\beta}(\lambda) = \sum_{n=-\infty}^{\infty} e^{-jn\lambda} E_{\mu,\nu} \left\{\left|\beta_{n,\mu,\nu}^{(k)}(Z^\ell)\right|^2\right\} \tag{34}$$

Then, the asumptotic coefficients $\{a_i\}$ of the vector $A_{M_0}(\mu,\nu)$ are given by the following expression:

$$S_{\mu,\nu,\beta}(\lambda) = \frac{1}{2\pi} \int_{-\pi}^{\pi} H(w)\ S_{\mu,\nu,\beta}(w,\lambda)\ dw \tag{35}$$

The error $h(\mu,\nu,k,g_{\ell,\mu})$ in lemma 6 takes then the following form:

$$h(\mu,\nu,k,g_{\ell,\mu}) = E_\mu X_0^2 - \frac{1}{2\pi} \int_{-\pi}^{\pi} H^*(w)\, S_{\mu,\nu,\beta}(w)\, dw \qquad (36)$$

; where H(w) the Fourier transform of the filter $\{a_n\}$ given by (35).

As in section 3, the properties of $h(\mu,\nu,k,g_{\ell,\mu})$, expressed by lemmas 5 and 6, are also valid in the asymptotic case. Thus, if M is a concave and closed class of measures μ, the $\sup_{\mu\epsilon M} h(\mu,\nu,k,g_{\ell,\mu})$ exists, it is unique, it is realized by some μ^* in M, and μ^* is such that:

$$h(\mu^*,\nu,k,g_{\ell,\mu^*}) = \sup_{\mu\epsilon M} h(\mu,\nu,k,g_{\ell,\mu})$$

Also, the conclusions in corollary 2 are still valid in the asymptotic case.

We will conclude this section with the consideration of the special classes of information processes that were studied in lemmas 3 and 4, in section 4. We will express our conclusions in a lemma whose proof is in appendix C.

Lemma 7

Let the noise channel $[R,\nu,R]$ be either deterministic, or Gaussian, stationary, additive, and with measure independent of the measure of the information process. Let M be either one of the following two convex and closed families of stationary processes:

1. M_1 is the class of processes with identical means and spectral densities.
2. M_2 is the class of processes with identical means and with spectral densities lying within a convex and closed family S of spectra.

For both the above families M, the measure μ^* in lemma 6 is the Gaussian measure, with spectral density uniquely determined for the class in 2.

Comparing the result in lemma 7 with the results in lemmas 3 and 4, we see that for both the considered families M, the two lower bounds represented by theorem 3 and lemma 6 are evaluated at the Gaussian measure. Furthermore, the first bound is exactly equal to the mean square error induced by the Gaussian measure on F. Then, the properties of qualitative robustness [10], [11], [12] are in general violated if the filters or smoothers that correspond to the above lower bounds are selected. This is so, because boundness and asymptotic continuity (as in [10]) are not in general satisfied then.

6. The Asymptotic Lower Bounds for A Class M

In this section we will evaluate the lower bounds of theorem 3 and lemma 6, for the first class M_1 of information process in lemma 7, and for the same noise channel specifications as in the lemma. We will consider the asymptotic case, and for the bound in lemma 6 we will assume arbitrary sliding step k, and finite arbitrary length ℓ of the sliding block window. We will assume zero mean information and noise processes.

Let us denote by $f_\nu(w)$ and $f_\mu(w)$ the spectral densities of the noise and each information process μ in M_1, respectively. Let us also denote:

$$R_{\mu\ell}(m) = E_\mu \left\{ X^o_{1-\ell} \; X^{T\,mk}_{mk+1-\ell} \right\} \tag{37}$$

$$R_{\nu\ell}(m) = E_\nu \left\{ Y^o_{1-\ell} \; Y^{T\,mk}_{mk+1-\ell} \right\} \tag{38}$$

The expression in (37) is not μ-dependent within the class M_1. Considering the asymptotic case with infinite past and future, and due to lemma 3, we have a well known expression for the lower bound in theorem 3. Indeed, if μ^* is the Gaussian measure within the class M_1, this lower bound is given by the following expression:

$$e(\mu^*,\nu,g_\infty^*) = \sup_{\mu\varepsilon M} e(\mu,\nu,\alpha_{\mu,\nu,\infty}) = \frac{1}{2\pi}\int_{-\pi}^{\pi} f_\mu(w)\,dw - \frac{1}{2\pi}\int_{-\pi}^{\pi}\frac{f_\mu^2(w)}{f_\mu(w)+f_\nu(w)}\,dw \quad (39)$$

We now proceed with the evaluation of the lower bound in lemma 7, for the class M_1. From the proof of lemma 7, in appendix C, we have:

$$\gamma^{(k)}_{m,\mu,\nu}(z^\ell) = E_\mu\left\{X_0\ X^T_{mk+1-\ell}{}_{mk}\right\}[R_{\mu\ell}(0) + R_{\nu\ell}(0)]^{-1}\ z^\ell\ ;\ \forall\ \mu\varepsilon M_1 \quad (40)$$

; where in M_1, $\gamma^{(k)}_{m,\mu,\nu}(z^\ell)$ is not μ-dependent.

Also, for the Gaussian measure μ^*, we have for $\beta^{(k)}_{m,\mu,\nu}(z^\ell)$ defined by (3):

$$\beta^{(k)}_{m,\mu^*,\nu}(z^\ell) \triangleq \gamma^{(k)}_{m,\mu,\nu}(z^\ell)\ ;\ \forall\ m \quad (41)$$

Using expressions (40) and (41), we have then:

$$E_{\mu,\nu}\left\{\gamma^{(k)}_{n,\mu,\nu}(Z^\ell)\cdot\gamma^{*(k)}_{m,\mu,\nu}(Z^\ell)\right\} = E_{\mu^*,\nu}\left\{\beta^{(k)}_{n,\mu^*,\nu}(Z^\ell)\ \beta^{*(k)}_{m,\mu^*,\nu}(Z^\ell)\right\} =$$
$$= E_\mu\left\{X_0\ X^T_{nk+1-\ell}{}_{nk}\right\}[R_{\mu\ell}(0) + R_{\nu\ell}(0)]^{-1}\ E_\mu\left\{X_0\ X^T_{mk+1-\ell}{}_{mk}\right\} \quad (42)$$

Let $\{r_i(\mu)\}$ be the autocovariance coefficients determined by the spectral density $f_\mu(w)$, within the class M_1. Those coefficients are μ-independent within the class M_1. We then obtain:

$$\sum_{n=-\infty}^{\infty} e^{-jnx}\ E_\mu\left\{X_0\ X^T_{nk+1-\ell}{}_{nk}\right\} = \left[\sum_{n=-\infty}^{\infty} e^{-jnx}\ r_{nk}(\mu),\ldots,\sum_{n=-\infty}^{\infty} e^{-jnx}\ r_{nk-\ell+1}(\mu)\right] \quad (43)$$

But after some manipulations, we obtain:

$$\sum_{n=-\infty}^{\infty} e^{-jnx}\ r_{nk-m}(\mu) = \frac{1}{k}\sum_{i=0}^{k-1} e^{-j\frac{x+2\pi i}{k}m}\ f_\mu\left(\frac{x+2\pi i}{k}\right)\ ;\ 0 \le m \le \ell-1 \quad (44)$$

Substituting expression (44) in expression (43), we also obtain:

$$\sum_{n=-\infty}^{\infty} e^{-jnx} E_\mu \left\{ X_0 \, X^T_{nk+1-\ell} \right\} = \frac{1}{k} \sum_{i=0}^{k-1} f_\mu\left(\frac{x+2\pi i}{k}\right) [1,\ldots,e^{-j\frac{x+2\pi i}{k}m},\ldots,e^{-j\frac{x+2\pi i}{k}(\ell-1)}] \tag{45}$$

We now substitute expressions (45) and (42) in expression (33), to obtain in a straightforward manner:

$$S_{\mu^*,\nu,\beta}(\lambda,x) = \frac{1}{k^2} \sum_{i=0}^{k-1} \sum_{n=0}^{k-1} f_\mu\left(\frac{x+2\pi i}{k}\right) f_\mu\left(\frac{\lambda+2\pi n}{k}\right) \cdot$$

$$\cdot [1,\ldots,e^{-j\frac{x+2\pi i}{k}m},\ldots,e^{-j\frac{x+2\pi i}{k}(\ell-1)}] \, [R_{\mu\ell}(0) + R_{\nu\ell}(0)]^{-1} \cdot$$

$$\cdot [1,\ldots,e^{j\frac{\lambda+2\pi n}{k}m},\ldots,e^{j\frac{\lambda+2\pi n}{k}(\ell-1)}]^T \tag{46}$$

For arbitrary finite ℓ and some k such that $1 \leq k \leq \ell$, the lower bound in lemma 6 will be given in this case by the expression:

$$h(\mu^*,\nu,k,g_{\ell,\mu^*}) = \frac{1}{2\pi} \int_{-\pi}^{\pi} f_\mu(w)\,dw - (2\pi)^{-2} \iint_{-\pi}^{\pi} H(x)\, S_{\mu^*,\nu,\beta}(\lambda,x)\, H^*(\lambda)\, dx\, d\lambda \tag{47}$$

; where $S_{\mu^*,\nu,\beta}(\lambda,x)$ is given by (46), and H(x) is the solution of (35) with $S_{\mu,\nu,\beta}(\lambda)$ determined by $\beta^{(k)}_{m,\mu,\nu}(z^\ell) = \gamma^{(k)}_{m,\mu,\nu}(z^\ell)$.

It is easy to see that for $k = \ell = 1$ the bound $h(\mu^*,\nu,1,g_{1,\mu^*})$ is the minimum error $E_{\mu,\nu} \left| X_0 - \sum_{i=-\infty}^{\infty} a_i Z_i \right|^2$ with respect to the filter $\{a_i\}$. Therefore, $h(\mu^*,\nu,1,g_{1,\mu^*})$ is then identical to the global lower bound $e(\mu^*,\nu,g^*_\infty)$ in (59). This was expected since both bounds are realized at the Gaussian measure μ^*, and for this measure the error minimization on $F(1,1)$ is identical to the error minimization on F and $F(L)$.

As in lemma 7, the bound $h(\mu^*,\nu,k,g_{\ell,\mu^*})$ in (47) is a strictly concave function of $f_\mu(w)$, for all k and ℓ. Therefore, if the class M_2 in lemma 7 is considered rather than the class M_1, there will be a unique spectral density f^*_μ, which will maximize the bound $h(\mu^*,\nu,k,g_{\ell,\mu^*})$ within the convex and closed family S of spectral densities. This maximum value will be the lower bound of lemma 7 in M_2, then. Similar argument holds for the global lower bound $e(\mu^*,\nu,g^*_\infty)$ of (39), in M_2.

7. Conclusions

Given some convex and closed family M of information processes, we developed two classes of lower bounds on the performance of filters and smoothers in M. The mean square performance criterion was adopted. The developed bounds are functions of conditional expectations of the type expressed by (3). Thus, they are, in general, functions of high order statistics in M. In other words, these bounds are, in general, tightly associated with the family M, in measure rather than in spectral characteristics only. So, if the class M is, for example, described by the linear contamination of a nominal measure, and if the Gaussian measure is not included in M, then the developed lower bounds will be nonlinear. Therefore, the adoption of linear filtering (as in [13]) will, in general, result in inferior performance across M then.

References

[1] P. J. Huber (1979), "Robust Smoothing," in Robustness in Statistics edited by R. L. Launer and G. Wilkinson, Academic Press.

[2] A. Jazwinski (1970), Stochastic Processes and Filtering Theory, Academic Press, New York, N.Y.

[3] J. W. Tukey (1977), Exploratory Data Analysis, Addison-Wesley, Reading, MA.

[4] P. J. Huber (1981), Robust Statistics, Joh Wiley, New York, N.Y.

[5] R. D. Martin (1979), "Robust Estimation of Time Series Autoregressions," in Robustness in Statistics edited by R. L. Launer and G. Wilkinson, Academic Press.

[6] R. D. Martin (1979), "Robust Estimation of Autoregressive Models," in Time Series Analysis, Surveys and Recent Developments, Inst. of Math. Stat.

[7] R. D. Martin, "Approximate Conditional-Mean Type Smoothers and Interpolators," Invited talk at the Heidelberg Workshop on Smoothing Techniques for Curve Estimation, Heidelberg, Germany, April 2-4, 1979.

[8] B. W. Stuck (1976), "Minimum Error Dispersion Linear Filtering of Symmetric Stable Processes," IEEE Trans. on Auto. Control, 17, 507-509.

[9] P. Velleman (1975), "Robust Non-linear Data Smoothing," Ph.D. Thesis, Dept. of Statistics, Princeton University.

[10] P. Papantoni-Kazakos and R. M. Gray (1979), "Robustness of Estimation on Stationary Observations," The Annals of Probability, Vol. 7, No. 6, 989-1002.

[11] P. Papantoni-Kazakos (1981), "Sliding Block Encoders that Are Rho-Bar Continuous Functions of their Input," IEEE Trans. Inf. Th. IT-27, 372-376.

[12] P. Papantoni-Kazakos, "A Game Theoretic Approach to Robust Filtering," University of Connecticut, EECS Dept., Technical Report TR-81-12, October 1981.

[13] Y. Hosoya (1978), "Robust Linear Extrapolation of Second Order Stationary Processes," The Annals of Probability, Vol. 6, 574-584.

[14] B. Kleiner, R. D. Martin, and D. J. Thomson (1979), "Robust Estimation of Power Spectra," Jour. Royal Stat. Soc. B, 41, No. 3.

[15] C. L. Mallows (1979), "Some Theory of Nonlinear Smoothers," Bell Labs report. To appear in the Annals of Statistics.

[16] H. V. Poor (1980), "On Robust Wiener Filtering," IEEE Trans. on Auto. Control, 25, 531-536.

Appendix A

Proof of Lemma 1

In the expression of $e(\mu,\nu,k,g_\ell)$ in (14), the term $E_{\mu,\nu}\, X_0^2$ is linear with respect to μ. Thus, we must only show that the expression

$$K(\mu,\nu,g_\ell) \triangleq E_{\mu,\nu}\left\{X_0\ V^T_{g_\ell}(Z)\right\} R^{-1}(\mu,\nu,g_\ell)\ E_{\mu,\nu}\left\{X_0\ V_{g_\ell}(Z)\right\} \tag{A.1}$$

is convex in M.

Let $\mu_1 \varepsilon M$ and $\mu_2 \varepsilon M$. Then, since M is convex and closed, for any $\varepsilon : 0 < \varepsilon < 1$ we also have $[\varepsilon\mu_1 + (1-\varepsilon)\mu_2] \varepsilon M$.

To simplify the notation, let us denote:

$$\begin{aligned} E_i &\triangleq E_{\mu_i,\nu}\left\{X_0\ V_{g_\ell}(Z)\right\} \ ;\ i = 1,2 \\ R_i &\triangleq R(\mu_i,\nu,g_\ell)\ ;\ i = 1,2 \end{aligned} \tag{A.2}$$

Then, it is clear that:

$$\begin{aligned} E_{\varepsilon\mu_1+(1-\varepsilon)\mu_2,\nu}\left\{X_0\ V_{g_\ell}(Z)\right\} &= \varepsilon\ E_1 + (1-\varepsilon)\ E_2 \\ R(\varepsilon\mu_1+(1-\varepsilon)\mu_2,\nu,g_\ell) &= \varepsilon\ R_1 + (1-\varepsilon)\ R_2 \end{aligned} \tag{A.3}$$

Substituting (A.3) in (A.1), we find:

$$K(\varepsilon\mu_1 + (1-\varepsilon)\ \mu_2,\nu,g_\ell) = [\varepsilon\ E_1 + (1-\varepsilon)\ E_2]^T\ [\varepsilon\ R_1 + (1-\varepsilon)\ R_2]^{-1}\ [\varepsilon\ E_1 + (1-\varepsilon)\ E_2] \tag{A.4}$$

For convexity, we want to prove that:

$$\frac{\partial^2}{\partial\varepsilon^2} K(\varepsilon\ \mu_1 + (1-\varepsilon)\ \mu_2,\nu,g_\ell)\Big|_{\varepsilon = 0} \geq 0 \tag{A.5}$$

We will need the following known equation:

$$\frac{\partial}{\partial x} M^{-1}(x) = - M^{-1}(x) \left[\frac{\partial}{\partial x} M(x)\right] M^{-1}(x) \tag{A.6}$$

; where M(x) a nonsingular square matrix, and x a scalar variable appearing in the components of the matrix M(x).

Differentiating the expression (A.4) twice with respect to ε, and through the application of expression (A.6), we find after some manipulations:

$$\frac{\partial^2}{\partial \varepsilon^2} K(\varepsilon\, \mu_1 + (1-\varepsilon)\, \mu_2, \nu, g_\ell)\Big|_{\varepsilon = 0} = 2[E_1 - R_1 R_2^{-1} E_2]^T R_2^{-1} [E_1 - R_1 R_2^{-1} E_2] \tag{A.7}$$

The expression (A.7) is clearly nonnegative since R_2^{-1} is a nonnegative definite matrix (correlation matrix). Thus, (A.5) is satisfied and the proof is complete.

Proof of Theorem 1

Due to lemma 1, there exists some μ^* in M that satisfies the supremum $\sup\limits_{\mu\varepsilon M} e(\mu,\nu,k,g_\ell)$. Even if μ^* is not unique, for every ε : $0 < \varepsilon < 1$ and every $\mu \varepsilon M$ we have that $[(1-\varepsilon)\, \mu^* + \varepsilon\, \mu]\ \varepsilon\ M$. This is due to the convexity and closeness of the set M. Also, even when there are more than one measures in M that satisfy the supremum $\sup\limits_{\mu\varepsilon M} e(\mu,\nu,k,g_\ell)$, there exists among them at least one μ^*, such that:

$$\frac{\partial}{\partial \varepsilon} e((1-\varepsilon)\mu^* + \varepsilon\, \mu,\nu,k,g_\ell)\Big|_{\varepsilon = 0} \leq 0 \ ; \ \forall\ \mu \varepsilon M \tag{A.8}$$

Due to (16), and using the expression in (14), we easily obtain:

$$\frac{\partial}{\partial \varepsilon} e(1-\varepsilon)\mu^* + \varepsilon\, \mu,\nu,k,g_\ell)\Big|_{\varepsilon = 0} = - e(\mu^*,\nu,k,g_\ell) + e(\mu,\nu,k,g_\ell,\{a_i^*\}) \tag{A.9}$$

; where $\{a_i^*\} = A^*_{\mu^*,\nu,g_\ell}$, and $A^*_{\mu^*,\nu,g_\ell}$ given by (15).

From (A.8) and (A.9) we conclude then:

There exists some μ^* in M that satisfies the supremum in lemma 1 and

is such that:

$$e(\mu,\nu,k,g_{\ell},\{a_i^*\}) \leq e(\mu^*,\nu,k,g_{\ell},\{a_i^*\}) \quad ; \ \forall \ \mu \varepsilon M$$

The part

$$e(\mu^*,\nu,k,g_{\ell},\{a_i^*\}) \leq e(\mu^*,\nu,k,g_{\ell},\{a_i\}) \quad ; \ \forall \ \{a_i\} \ \varepsilon \ F(k,g_{\ell})$$

is trivially true since $\{a_i^*\}$ satisfies the infimum in (16) at the process μ^*. The proof of the theorem is now complete.

Appendix B

Proof of Lemma 2

Let μ_1,μ_2 belong to M. Let $\varepsilon : 0 < \varepsilon < 1$. Then, since M is convex and closed we also have $[\varepsilon\,\mu_1 + (1-\varepsilon)\mu_2] \in M$. Using the symbol $f_{\mu,\nu}(z^{j-i})$ for the $(j-i)$-dimensional density function induced by μ and ν, as introduced in section 2, we easily find:

$$\alpha_{\varepsilon\mu_1+(1-\varepsilon)\mu_2,\nu,(i,j)}(z^{j-i}) = \frac{\varepsilon f_{\mu_1,\nu}(z^{j-i})\alpha_{\mu_1,\nu,(i,j)}(z^{j-i})+(1-\varepsilon)f_{\mu_2,\nu}(z^{j-i})\alpha_{\mu_2,\nu,(i,j)}(z^{j-i})}{\varepsilon f_{\mu_1,\nu}(z^{j-i})+(1-\varepsilon)f_{\mu_2,\nu}(z^{j-i})} \tag{B.1}$$

Directly from (B.1) we also find:

$$\frac{\partial}{\partial\varepsilon}\,\alpha_{\varepsilon\mu_1+(1-\varepsilon)\mu_2,\nu,(i,j)}(z^{j-i}) = \frac{f_{\mu_1,\nu}(z^{j-i})f_{\mu_2,\nu}(z^{j-i})[\alpha_{\mu_1,\nu,(i,j)}(z^{j-i}) - \alpha_{\mu_2,\nu,(i,j)}(z^{j-i})]}{[\varepsilon f_{\mu_1,\nu}(z^{j-i})+(1-\varepsilon)f_{\mu_2,\nu}(z^{j-i})]^2} \tag{B.2}$$

$$\frac{\partial^2}{\partial\varepsilon^2}\,\alpha_{\varepsilon\mu_1+(1-\varepsilon)\mu_2,\nu,(i,j)}(z^{j-i}) = -2\left[f_{\mu_1,\nu}(z^{j-i})-f_{\mu_2,\nu}(z^{j-i})\right]f_{\mu_1,\nu}(z^{j-i})f_{\mu_2,\nu}(z^{j-i})\left[\alpha_{\mu_1,\nu,(i,j)}(z^{j-i}) - \alpha_{\mu_2,\nu,(i,j)}(z^{j-i})\right]\left[\varepsilon f_{\mu_1,\nu}(z^{j-i})+(1-\varepsilon)f_{\mu_2,\nu}(z^{j-i})\right]^{-2} \tag{B.3}$$

$$\frac{\partial}{\partial\varepsilon}\alpha_{\varepsilon\mu_1+(1-\varepsilon)\mu_2,\nu,(i,j)}(z^{j-i})\Big|_{\varepsilon=0} = \frac{f_{\mu_1,\nu}(z^{j-i})}{f_{\mu_2,\nu}(z^{j-i})}\left[\alpha_{\mu_1,\nu,(i,j)}(z^{j-i}) - \alpha_{\mu_2,\nu,(i,j)}(z^{j-i})\right] \tag{B.4}$$

$$\frac{\partial^2}{\partial\varepsilon^2}\alpha_{\varepsilon\mu_1+(1-\varepsilon)\mu_2,\nu,(i,j)}(z^{j-i})\Big|_{\varepsilon=0} = 2\,\frac{f_{\mu_1,\nu}(z^{j-i})}{f_{\mu_2,\nu}(z^{j-i})}\left[1-\frac{f_{\mu_1,\nu}(z^{j-i})}{f_{\mu_2,\nu}(z^{j-i})}\right]\cdot \left[\alpha_{\mu_1,\nu,(i,j)}(z^{j-i}) - \alpha_{\mu_2,\nu,(i,j)}(z^{j-i})\right] \tag{B.5}$$

$$\begin{aligned}\frac{\partial^2}{\partial\varepsilon^2}E_{\varepsilon\mu_1+(1-\varepsilon)\mu_2,\nu}\left\{\alpha^2_{\varepsilon\mu_1+(1-\varepsilon)\mu_2,\nu,(i,j)}(z^{j-i})\right\} &= 4\,E_{\mu_1-\mu_2,\nu}\left\{\alpha_{\varepsilon\mu_1+(1-\varepsilon)\mu_2,\nu,(i,j)}(z^{j-i})\cdot\frac{\partial}{\partial\varepsilon}\alpha_{\varepsilon\mu_1+(1-\varepsilon)\mu_2,\nu,(i,j)}(z^{j-i})\right\}\\ &+ 2\,E_{\varepsilon\mu_1+(1-\varepsilon)\mu_2,\nu}\left\{\left[\frac{\partial}{\partial\varepsilon}\alpha_{\varepsilon\mu_1+(1-\varepsilon)\mu_2,\nu,(i,j)}(z^{j-i})\right]^2\right.\\ &\left.+\alpha_{\varepsilon\mu_1+(1-\varepsilon)\mu_2,\nu,(i,j)}(z^{j-i})\cdot\frac{\partial^2}{\partial\varepsilon^2}\alpha_{\varepsilon\mu_1+(1-\varepsilon)\mu_2,\nu,(i,j)}(z^{j-i})\right\}\end{aligned} \tag{B.6}$$

Applying expressions (B.2) - (B.5) in expression (B.6), we finally find:

$$\begin{aligned}&\frac{\partial^2}{\partial\varepsilon^2}E_{\varepsilon\mu_1+(1-\varepsilon)\mu_2,\nu}\left\{\alpha^2_{\varepsilon\mu_1+(1-\varepsilon)\mu_2,\nu,(i,j)}(z^{j-i})\right\}\Big|_{\varepsilon=0} =\\ &= 2\,E_{\mu_1}\left\{\frac{f_{\mu_1,\nu}(z^{j-i})}{f_{\mu_2,\nu}(z^{j-i})}\left[\alpha_{\mu_1,\nu,(i,j)}(z^{j-i}) - \alpha_{\mu_2,\nu,(i,j)}(z^{j-i})\right]^2\right\}\end{aligned} \tag{B.7}$$

The expression in (B.7) is nonnegative. It is strictly positive iff $\alpha_{\mu_1,\nu,(i,j)}(z^{j-i})$ and $\alpha_{\mu_2,\nu,(i,j)}(z^{j-i})$ are not equal almost everywhere in R^{j-i}. Thus, $E_{\mu,\nu}\left\{\alpha^2_{\mu,\nu,(i,j)}(Z^{j-i})\right\}$ is convex in M. Since $E_\mu X_0^2$ is linear in M, we conclude that $e(\mu,\nu,\alpha_{\mu,\nu,(i,j)})$ is concave in M.

Proof of Theorem 3

According to lemma 2, there exists some μ^* in M that realizes the supremum of $e(\mu,\nu,\alpha_{\mu,\nu,(i,j)})$ in M. Even if there are more than one such measures, there exists at least one such that:

For every $\varepsilon : 0 < \varepsilon < 1$, and every μ in M:

$$\frac{\partial}{\partial\varepsilon} e\left((1-\varepsilon)\mu^*+\varepsilon\mu,\nu,\alpha_{(1-\varepsilon)\mu^*+\varepsilon\mu,\nu,(i,j)}\right)\Big|_{\varepsilon=0} \leq 0 \;;\; \forall\, \mu\varepsilon M \tag{B.8}$$

Notice that since M is convex and closed, $[(1-\varepsilon)\mu^*+\varepsilon\mu]\ \varepsilon\ M$. But from (24) we have:

$$e\left((1-\varepsilon)\mu^*+\varepsilon\mu,\nu,\alpha_{(1-\varepsilon)\mu^*+\varepsilon\mu,\nu,(i,j)}\right) = E_{(1-\varepsilon)\mu^*+\varepsilon\mu} X_0^2 - E_{(1-\varepsilon)\mu^*+\varepsilon\mu,\nu}\left\{\alpha^2_{(1-\varepsilon)\mu^*+\varepsilon\mu,\nu,(i,j)}(Z \tag{B.9}$$

Applying then the results in (B.1)-(B.5) in (B.9), we find:

$$\frac{\partial}{\partial\varepsilon} e\left((1-\varepsilon)\mu^*+\varepsilon\mu,\nu,\alpha_{(1-\varepsilon)\mu^*+\varepsilon\mu,\nu,(i,j)}\right)\Big|_{\varepsilon=0} =$$

$$= -\left[E_{\mu^*} X_0^2 - E_{\mu^*,\nu}\left\{\alpha^2_{\mu^*,\nu,(i,j)}(Z^{j-i})\right\}\right] + E_\mu X_0^2$$

$$- E_{\mu,\nu}\left\{\alpha^2_{\mu^*,\nu,(i,j)}(Z^{j-i})\right\} -$$

$$- 2\, E_{\mu^*,\nu}\left\{\alpha_{\mu^*,\nu,(i,j)}(Z^{j-i})\, \frac{f_{\mu,\nu}(Z^{j-i})\left[\alpha_{\mu,\nu,(i,j)}(Z^{j-i}) - \alpha_{\mu^*,\nu,(i,j)}(Z^{j-i})\right]}{f_{\mu^*,\nu}(Z^{j-i})}\right\} \tag{B.10}$$

But,

$$E_{\mu^*,\nu}\left\{\alpha_{\mu^*,\nu,(i,j)}(Z^{j-i})\frac{f_{\mu,\nu}(Z^{j-i})\left[\alpha_{\mu,\nu,(i,j)}(Z^{j-i})-\alpha_{\mu^*,\nu,(i,j)}(Z^{j-i})\right]}{f_{\mu^*,\nu}(Z^{j-i})}\right\}=$$

$$=\int_{R^{j-i}} d\,z^{j-i}\, f_{\mu^*,\nu}(z^{j-i})\,\alpha_{\mu^*,\nu,(i,j)}(z^{j-i})\frac{f_{\mu,\nu}(z^{j-i})\left[\alpha_{\mu,\nu,(i,j)}(z^{j-i})-\alpha_{\mu^*,\nu,(i,j)}(z^{j-i})\right]}{f_{\mu^*,\nu}(z^{j-i})}=$$

$$=E_{\mu,\nu}\left\{\alpha_{\mu,\nu,(i,j)}(Z^{j-i})\,\alpha_{\mu^*,\nu,(i,j)}(Z^{j-i})\right\}-E_{\mu,\nu}\left\{\alpha^2_{\mu^*,\nu,(i,j)}(Z^{j-i})\right\} \tag{B.11}$$

Also,

$$e(\mu^*,\nu,\ \alpha_{\mu^*,\nu,(i,j)}) = E_{\mu^*}X_0^2 - E_{\mu^*,\nu}\left\{\alpha^2_{\mu^*,\nu,(i,j)}(Z^{j-i})\right\} \tag{B.12}$$

$$e(\mu,\nu,\ \alpha_{\mu^*,\nu,(i,j)}) = E_{\mu,\nu}\left\{X_0 - \alpha_{\mu^*,\nu,(i,j)}(Z_i^j)\right\}^2 =$$

$$= E_\mu X_0^2 - 2\,E_{\mu,\nu}\left\{\alpha_{\mu,\nu,(i,j)}(Z^{j-i})\,\alpha_{\mu^*,\nu,(i,j)}(Z^{j-i})\right\} + E_{\mu,\nu}\left\{\alpha^2_{\mu^*,\nu,(i,j)}(Z^{j-i})\right\} \tag{B.13}$$

Substituting (B.11), (B.12), and (B.13), in (B.10), we obtain:

$$\frac{\partial}{\partial\varepsilon}\, e\Big((1-\varepsilon)\mu^*+\varepsilon\,\mu,\nu,\ \alpha_{(1-\varepsilon)\mu^*+\varepsilon\,\mu,\nu,(i,j)}\Big)\Big|_{\varepsilon\,=\,0} = e(\mu,\nu,\ \alpha_{\mu^*,\nu,(i,j)}) -$$

$$- e(\mu^*,\nu,\ \alpha_{\mu^*,\nu,(i,j)}) \tag{B.14}$$

Finally, from (b.8) and (B.14) we obtain:

$$e(\mu,\nu,\ \alpha_{\mu^*,\nu,(i,j)}) \le e(\mu^*,\nu,\ \alpha_{\mu^*,\nu,(i,j)})\ ;\ \forall\ \mu\varepsilon M \tag{B.15}$$

Expression (B.15) proves the left part of the game. The right part is trivially satisfied due to (24).

Proof of Lemma 3

Consider the linear class $\mathcal{F}(L)$ of filters and smoothers, as introduced in section 2. Let us denote by $e_L(\mu,\nu)$ the infimum of the error in (6) in $\mathcal{F}(L)$, at the neasure μ. Then, due to the fact that $e_L(\mu,\nu)$ is only a function of first and second order statistics, we have that for any μ_1, μ_2 in the class $\mathcal{M}$ of the lemma:

$$e_L(\mu_1,\nu) = e_L(\mu_2,\nu) = e_L(\mu^*,\nu) \ ; \ \forall\ \mu_1,\mu_2 \ \varepsilon\ \mathcal{M} \qquad \text{(B.16)}$$

; where by μ^* we have denoted the Gaussian measure in $\mathcal{M}$.

Bue due to (9) we have:

$$e_L(\mu,\nu) \geq e(\mu,\nu,\ \alpha_{\mu,\nu}) \ ; \ \forall\ \mu\varepsilon\mathcal{M} \qquad \text{(B.17)}$$

Also, for the Gaussian measure μ^* we have:

$$e_L(\mu^*,\nu) = e(\mu^*,\nu,\ \alpha_{\mu^*,\nu}) \qquad \text{(B.18)}$$

; where (B.18) is true since μ^* and ν jointly Gaussian here.

Thus, from (B.16), (B.17) and (B.18), we obtain:

$$e_L(\mu^*,\nu) = e(\mu^*,\nu,\ \alpha_{\mu^*,\nu}) \geq e(\mu,\nu,\ \alpha_{\mu,\nu}) \ ; \ \forall\ \mu\varepsilon\mathcal{M}$$

Thus, it is clear that the pair $(\mu^*,\ \alpha_{\mu^*,\nu})$ is a solution of the game in theorem 3. The solution is unique if for any μ_1,μ_2 in $\mathcal{M}$ the conditional expectations $\alpha_{\mu_1,\nu}(z)$ and $\alpha_{\mu_2,\nu}(z)$ can not be equal almost everywhere.

Appendix C

Proof of Lemma 5

For any g_ℓ, the infimum $e(\mu,\nu,k,g_\ell)$ is given by expression (14). We see in a relatively straightforward manner, that the following equation holds ; where A^*_{μ,ν,g_ℓ} is given by (15):

$$E_{\mu,\nu}\left\{X_0\, V^T_{g_\ell}(Z)\right\} R^{-1}(\mu,\nu,g_\ell)\, E_{\mu,\nu}\left\{X_0\, V_{g_\ell}(Z)\right\} = E_{\mu,\nu}\left\{g_\ell(Z^\ell)\, A^{*T}_{\mu,\nu,g_\ell}\, B_{\mu,\nu,(k),-n,m-n-1}(Z^\ell)\right\} \tag{C.1}$$

Let A be an arbitrary vector. Then, substituting A^*_{μ,ν,g_ℓ} by A in (C.1) and applying the Schwartz inequality, we obtain:

$$E_{\mu,\nu}\left\{g_\ell(Z^\ell) A^T B_{\mu,\nu,(k),-n,m-n-1}(Z^\ell)\right\} \le E_{\mu,\nu}\left\{\left|g_\ell(Z^\ell)\right|\left|A^T B_{\mu,\nu,(k),-n,m-n-1}(Z^\ell)\right|\right\}$$

$$\le E^{1/2}_{\mu,\nu}\left\{\left|g_\ell(Z^\ell)\right|^2\right\}\cdot E^{1/2}_{\mu,\nu}\left\{\left|A^T B_{\mu,\nu,(k)-n,m-n-1}(Z^\ell)\right|^2\right\} =$$

$$= E^{1/2}_{\mu,\nu}\left\{\left|g_\ell(Z^\ell)\right|^2\right\}\cdot [A^T M_0(\mu,\nu) A]^{1/2} \tag{C.2}$$

with equality everywhere in (C.2) iff:

$$g_\ell(z^\ell) = C\, A^T B_{\mu,\nu,(k),-n,m-n-1}(z^\ell) \tag{C.3}$$

For $g_\ell(z^\ell)$ as in (C.3), the upper bound in (C.2) is maximized for vector A such that it minimizes the mean square error expression $E_{\mu,\nu}\left\{X_0 - A^T B_{\mu,\nu,(k),-n,m-n-1}(Z^\ell)\right\}^2$. The proof is now complete.

Proof of Lemma 6

We will present the proof in two steps.

i) Consider the matrix $M_0(\mu,\nu)$ in (30). Let A be some arbitrary m-dimensional constant vector. We will show that the quadratic form

$A^T M_0(\mu,\nu)A$ is convex in μ on M. We will refer to the proof of lemma 2, in appendix B. Let $\mu_1,\mu_2 \in M$, let $\varepsilon : 0 < \varepsilon < 1$. Then, $[(1-\varepsilon)\mu_2 + \varepsilon\, \mu_1] \varepsilon M$. It can be easily seen that each $\beta^{(k)}_{i,\varepsilon\mu_1+(1-\varepsilon)\mu_2,\nu}(z^{\ell})$ in the matrix $M_0(\varepsilon\, \mu_1+(1-\varepsilon)\mu_2,\nu)$ has the same form as in (B.1). Then, using the expressions (B.1)-(B.5) and differentiating the quadratic form $A^T M_0(\varepsilon\, \mu_1+(1-\varepsilon)\mu_2,\nu)A$ twice with respect to ε, we obtain:

$$\frac{\partial^2}{\partial\varepsilon^2} A^T M_0(\varepsilon\, \mu_1+(1-\varepsilon)\mu_2,\nu)\, A\Big|_{\varepsilon = 0} = 2\, E_{\mu_1,\nu}\left\{\sum_i a_i \left[\beta^{(k)}_{i,\mu_1,\nu}(z^{\ell}) - \beta_{i,\mu_2,\nu}(z^{\ell})\right]\right\}^2 \tag{C.4}$$

; where $\{a_i\}$ the components of the vector A.

The expression in (C.4) is strictly positive for nonzero A and notrivial measures. Thus, $A^T M_0(\mu,\nu)A$ is then strictly convex in μ on M.

ii) We will show that $A_{M_0}(\mu,\nu)M_0(\mu,\nu)A_{M_0}(\mu,\nu)$ is strictly convex in μ on M.

Since $E_\mu\, X_0^2$ is linear in μ, this will also show that $h(\mu,\nu,k,g_{\ell,\mu})$ is strictly concave in μ on M. Let $\mu_1,\mu_2 \in M$. Let $\varepsilon : 0 < \varepsilon < 1$. Then $[(1-\varepsilon)\mu_2+\varepsilon\, \mu_1] \in M$. Due to part i) of the proof, we have:

$$\lambda_{M_0}(\varepsilon\, \mu_1+(1-\varepsilon)\mu_2,\nu) \triangleq A^T_{M_0}(\varepsilon\, \mu_1+(1-\varepsilon)\mu_2,\nu)M_0(\varepsilon\, \mu_1+(1-\varepsilon)\mu_2,\nu)A_{M_0}(\varepsilon\, \mu_1+(1-\varepsilon)\mu_2,\nu)$$

$$< \varepsilon\, A^T_{M_0}(\varepsilon\, \mu_1+(1-\varepsilon)\mu_2,\nu)M_0(\mu_1,\nu)A_{M_0}(\varepsilon\, \mu_1+(1-\varepsilon)\mu_2,\nu) +$$

$$+ (1-\varepsilon)A^T_{M_0}(\varepsilon\, \mu_1+(1-\varepsilon)\mu_2,\nu)M_0(\mu_2,\nu)A_{M_0}(\varepsilon\, \mu_1+(1-\varepsilon)\mu_2,\nu) \tag{C.5}$$

But, by definition:

$$\begin{aligned} A^T_{M_0}(\varepsilon\, \mu_1+(1-\varepsilon)\mu_2,\nu)M_0(\mu_1,\nu)A_{M_0}(\varepsilon\, \mu_1+(1-\varepsilon)\mu_2,\nu) &\le \lambda_{M_0}(\mu_1,\nu) \\ A^T_{M_0}(\varepsilon\, \mu_1+(1-\varepsilon)\mu_2,\nu)M_0(\mu_2,\nu)A_{M_0}(\varepsilon\, \mu_1+(1-\varepsilon)\mu_2,\nu) &\le \lambda_{M_0}(\mu_2,\nu) \end{aligned} \tag{C.6}$$

From (C.5) and (C.6) we thus obtain:

$$\lambda_{M_0}(\varepsilon\,\mu_1+(1-\varepsilon)\mu_2,\nu) < \varepsilon\,\lambda_{M_0}(\mu_1,\nu) + (1-\varepsilon)\,\lambda_{M_0}(\mu_2,\nu) \tag{C.7}$$

; where $\lambda_{M_0}(\mu,\nu) \triangleq A^T_{M_0}(\mu,\nu)M_0(\mu,\nu)\,A_{M_0}(\mu,\nu)$

<u>Proof of Lemma 7</u>

i) Given μ, let us define:

$$\gamma^{(k)}_{m,\mu,\nu}(z^\ell) = E_\mu\left\{X_0\,X^{T\,mk}_{mk+1-\ell}\right\}[R_{\mu\ell}+R_{\nu\ell}]^{-1}\,z^\ell \tag{C.8}$$

$$R_{\mu\ell} = E_\mu\left\{X^{mk}_{mk+1-\ell}\,X^{T\,mk}_{mk+1-\ell}\right\} \tag{C.9}$$

$$R_{\nu\ell} = E_\nu\left\{Y^{mk}_{mk+1-\ell}\,Y^{T\,mk}_{mk+1-\ell}\right\} \tag{C.10}$$

The $\gamma^{(k)}_{m,\mu,\nu}(z^\ell)$ is the linear mean square estimate of X_0, given $z^{mk}_{mk+1-\ell}$. For the Gaussian measure μ^*, $\gamma^{(k)}_{m,\mu,\nu}(z^\ell)$ is exactly the same with the $\beta^{(k)}_{m,\mu,\nu}(z^\ell)$ in (3). If in the matrix $M_0(\mu,\nu)$ we substitute $\beta^{(k)}_{m,\mu,\nu}(z^\ell)$ by $\gamma^{(k)}_{m,\mu,\nu}(z^\ell)$, we will obtain a matrix $M_{0,\gamma}(\mu,\nu)$ with the same convexity properties as in $M_0(\mu,\nu)$ and with a maximum inner product $\lambda_{M_0,\gamma}(\mu,\nu)$ (denoted in the proof of lemma 6), that is in general lower than the inner product $\lambda_{M_0}(\mu,\nu)$ in (32). This is clear from the proofs of lemmas 5 and 6. Also, the components of the matrix $M_{0,\gamma}(\mu,\nu)$ are strictly functions of covariance coefficients (see (C.8)), thus the matrices $M_{0,\gamma}(\mu,\nu)$ are identical within the class M_1. Thus, also:

$$\lambda_{M_0,\gamma}(\mu_1,\nu) = \lambda_{M_0,\gamma}(\mu_2,\nu)\ ;\ \forall\ \mu_1,\mu_2 \in M_1 \tag{C.11}$$

Also, given μ in M_1 :

$$\lambda_{M_0,\gamma}(\mu,\nu) \leq \lambda_{M_0}(\mu,\nu) \tag{C.12}$$

If μ^* the Gaussian measure in M_1 we finally obtain combining (C.11) and (C.12).

$$\lambda_{M_0}(\mu^*,\nu) = \lambda_{M_0,\gamma}(\mu^*,\nu) \geq \lambda_{M_0,\gamma}(\mu,\nu) \;\; ; \; \forall \; \mu \varepsilon M_1 \tag{C.13}$$

This proves the lemma for the class M_1.

ii) As in part i) of the proof of lemma 6, for any vector A the quadratic form $A^T M_{0,\gamma}(\mu,\nu)A$ is strictly convex in μ. It is also a function of spectral characteristics only. Thus, it is strictly convex in f_μ on S; where f_μ the spectral density of the process μ. Following exactly the same procedure as in part ii) of the proof of lemma 6 we can then show that there exists a unique f^* in S such that

$$\lambda_{M_0,\gamma,f^*}(\mu,\nu) \leq \lambda_{M_0,\gamma,f}(\mu,\nu) \;\; ; \; \forall \; f \in S \tag{C.14}$$

; where $\lambda_{M_0,\gamma,f}(\mu,\nu)$ the inner product $\lambda_{M_0,\gamma}(\mu,\nu)$ if the process μ has spectral density f.

Thus the measure μ^* for the class M_2 of the lemma is the Gaussian with spectral density f^*, signified by (C.14).

The dimensionality of the observation sequence was never used in our proof. Thus, our results hold asymptotically, as well as nonasymptotically.

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered) UNCLASSIFIED

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER: AFOSR-TR- 83-0110
2. GOVT ACCESSION NO.:
3. RECIPIENT'S CATALOG NUMBER:
4. TITLE (and Subtitle): PERFORMANCE BOUNDS IN ROBUST FILTERING AND SMOOTHING
5. TYPE OF REPORT & PERIOD COVERED: Technical
6. PERFORMING ORG. REPORT NUMBER: TR 82-6
7. AUTHOR(s): P. Papantoni-Kazakos
8. CONTRACT OR GRANT NUMBER(s): AFOSR-78-3695
9. PERFORMING ORGANIZATION NAME AND ADDRESS: School of Engineering, University of Connecticut, Storrs, Connecticut 06268
10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS: 61102F 2304/A5
11. CONTROLLING OFFICE NAME AND ADDRESS: Mathematical and Information Sciences Directorate, Air Force Office of Scientific Research, Bolling AFB, DC 20332
12. REPORT DATE: 20 October 1982
13. NUMBER OF PAGES: 39
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office):
15. SECURITY CLASS. (of this report): Unclassified
15a. DECLASSIFICATION/DOWNGRADING SCHEDULE:
16. DISTRIBUTION STATEMENT (of this Report): Approved for public release; distribution unlimited.
17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report):
18. SUPPLEMENTARY NOTES:
19. KEY WORDS (Continue on reverse side if necessary and identify by block number):
20. ABSTRACT (Continue on reverse side if necessary and identify by block number):

Filtering and smoothing within a convex and closed family of stationary processes is considered. The performance criterion adopted is the mean square error. Using a saddle point game approach, we develop two classes of lower bounds for this error. Those bounds can be used as evaluation measure in the design of robust filters and smoothers.

DD FORM 1 JAN 73 1473

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

4 -

DT